

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 21, 2001, 18:18:09 ; Search time 7124 Seconds
(without alignments)

11995.973 Million cell updates/sec

Title: US-09-515-806-1

Perfect score: 5525

Sequence: 1 tgcgccacgcgtccgcacc.....aatgctttcatataacctgca 5525

Scoring table:

OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1344157 seqs, 7733874588 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:

- 1: gb_ba1.*
- 2: gb_ba2.*
- 3: gb_ba3.*
- 4: gb_in1.*
- 5: gb_in2.*
- 6: gb_in3.*
- 7: gb_om.*
- 8: gb_ov.*
- 9: gb_pat1.*
- 10: gb_pat2.*
- 11: gb_ph.*
- 12: gb_pl1.*
- 13: gb_pl2.*
- 14: gb_pl3.*
- 15: gb_pl4.*
- 16: em_ba1.*
- 17: em_ba2.*
- 18: em_fun.*
- 19: em_htgo_hum.*
- 20: em_htgo_inv.*
- 21: em_htgo_rnd.*
- 22: em_htgo_hum1.*
- 23: em_htgo_hum2.*
- 24: em_htgo_hum3.*
- 25: em_htgo_hum4.*
- 26: em_htgo_hum5.*
- 27: em_htgo_hum6.*
- 28: em_htgo_hum7.*
- 29: em_htgo_hum8.*
- 30: em_htgo_inv1.*
- 31: em_htgo_inv2.*
- 32: em_htgo_other.*
- 33: em_htgo_rnd.*
- 34: em_hum1.*
- 35: em_hum2.*
- 36: em_hum3.*
- 37: em_hum4.*
- 38: em_hum5.*
- 39: em_hum6.*
- 40: em_hum7.*
- 41: em_in.*
- 42: em_om.*
- 43: em_or.*

- 44: em_ov.*
- 45: em_pat.*
- 46: em_ph.*
- 47: em_pl.*
- 48: em_ro.*
- 49: em_sts.*
- 50: em_sy.*
- 51: em_un.*
- 52: em_v1.*
- 53: gb_sts1.*
- 54: gb_sts2.*
- 55: gb_sts3.*
- 56: gb_sy.*
- 57: gb_un.*
- 58: gb_v11.*
- 59: gb_v12.*
- 60: gb_htg1.*
- 61: gb_htg2.*
- 62: gb_htg3.*
- 63: gb_htg4.*
- 64: gb_htg5.*
- 65: gb_htg6.*
- 66: gb_htg7.*
- 67: gb_htg8.*
- 68: gb_htg9.*
- 69: gb_htg10.*
- 70: gb_htg11.*
- 71: gb_htg12.*
- 72: gb_htg13.*
- 73: gb_htg14.*
- 74: gb_htg15.*
- 75: gb_htg16.*
- 76: gb_htg17.*
- 77: gb_htg18.*
- 78: gb_htg19.*
- 79: gb_htg20.*
- 80: gb_htg21.*
- 81: gb_htg22.*
- 82: gb_htg23.*
- 83: gb_htg24.*
- 84: gb_htg25.*
- 85: gb_pr1.*
- 86: gb_pr2.*
- 87: gb_pr3.*
- 88: gb_pr4.*
- 89: gb_pr5.*
- 90: gb_pr6.*
- 91: gb_pr7.*
- 92: gb_pr8.*
- 93: gb_pr9.*
- 94: gb_ro1.*
- 95: gb_ro2.*
- 96: gb_in4.*
- 97: gb_pr10.*
- 98: em_ba3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	5107	92.4	5163	9	AX056426 Sequence
2	4878	88.3	4994	85	AB037759 Homo sapi
3	3332	60.3	3351	93	AL157497 Homo sapi
4	2650	48.0	2986	89	AK027011 Homo sapi
5	2111	38.2	2162	93	AJ243428 Homo sapi
6	1032	18.7	1138	93	AL137627 Homo sapi
7	919	16.6	1917	93	AL137676 Homo sapi
8	539	9.8	164297	62	AC012138 Homo sapi

Db 3306 ATGAGCACACGAGCTGCCCTATTATGACACACAGCGGGATGCTGGTGATGCTTCCTT 3365
Qy 3409 ttgacctggagatcccttttgcaagatatgtgcaagaataatattattgaatttaaac 3468
Db 3366 TTGACCTGGCGGATCCCTTTTGCAAGATATGTGCAAGAATAATATATTGAATTTAAAC 3425
Qy 3469 gatactgcatagaacgtgtgtttcagggccgcgaagttagatcgatttcacccaagaac 3528
Db 3426 GATACGATAGAACGTGTGTTCCAGCGCGCAAGTTAGATCGATTTCATCCCAAGAAC 3485
Qy 3529 ttctggagtgcatattgatattgtcaattctacccaacagcctttctgccactgtcg 3588
Db 3486 TTCTGGAGTGTGATTTGATATTGCTCACTTCTACCAACAGCTTCTGCCACTGTG 3545
Qy 3589 aaattatctacatctatgaataatcatcaagagtttccagcacttcaggaagaataatt 3648
Db 3546 AATATTATCATATCTATGAATAATCATCAAGAGTTTCCAGCATTTCCAGGAAGAAT 3605
Qy 3649 acagtatttttgaaccataccatgtattgaaagcaatactcttactctggggtac 3708
Db 3606 ACAGTATTTATTTGAACCATACCATCTTATTGAAAGCAATACTTTACACTGTGGATCC 3665
Qy 3709 cagaagaataaactcagtcagtcagtcattctctgtatgtatgctgtgacagagaagctga 3768
Db 3666 CAGAAGATAAATCACTCAAGTCTACATTTATCTGTATGATGCTGTGACAGAAAGCTGA 3725
Qy 3769 cgaggagagaagtggaagctaaattttgtaattctgtcttctcttaattgtctgtgc 3828
Db 3726 CGAGGAGAGAAGTGAAGCTAAATTTTGTAACTGCTTTTCTTAATATGCTGTGTC 3785
Qy 3829 gactctacaagtttattgaacagaagggagatttgcaagatctttatgccacaataaatt 3888
Db 3786 GACTCTACAAGTTTATTGAACAGAGGGAGATTGTCAGATCTTATGCCAACATAAAT 3845
Qy 3889 cattaataaacaagaacaggttatgcagctgggtgaggtatggcttaagaagacctag 3948
Db 3846 CATTAATAAACAGAAACAGGTTATGTCAGTTGGTGAAGTATGGCTTTAAAGACCTAG 3905
Qy 3949 aggaggtgtgtgactgttgaagaactcggcatacgaattcaggtcttgatcaattgg 4008
Db 3906 AGGAGGTGTGTGACTGTTGAAGAACTCGGCATCAAGTTACAGTCTTGATCAATTTGG 3965
Qy 4009 gcttgggtttacaaggtgcagcagcacaatggaaatcattccagctttgtggtttcatca 4068
Db 3966 GCTTGGTTTACAAGGTGCAGCAGCACAATGGAATCATCTTCCAGTTTGTGGCTTCATCA 4025
Qy 4069 aacgaagcaaaaggtgtactgtgaatcctgcagctggagggagagatgacctgtga 4128
Db 4026 AACGAAGCAAAAGGGCTGTACTCTGAAATCCTCGACGCTGGAGGCAGATATGACCTGTGA 4085
Qy 4129 ttcccagtttagagggccacaagctctgggcccagttcccactgpcattggggtcagca 4188
Db 4086 TTCCCAGTTTACAGGGCCACAGCTCTGGGCCAGTTTCCCAGTGCATTTGGGTTCAGCA 4145
Qy 4189 tagctatagaacaagatatctgtctgtctgtctgtctgtctgtctgtctgtctgtctgt 4248
Db 4146 TAGCTATAGACAAGATATCTGCTGCTGCTCCTCAACATGGAGGAATCTGTTACAATAAGCT 4205
Qy 4249 ctgtgaacctcctgt 4308
Db 4206 CTGTGACCTCTCTGT 4265
Qy 4309 cccagaactctggacagcaggtcatcacagcagagaatactgtacagctggtcacagctcc 4368
Db 4266 CCCAGAACTCTGGACAGCGGATCATCACAGCAATCATGTACGACTGGTCACAGTCCC 4325
Qy 4369 aagagaattacaagagttactgcagacatcatgaaatcaccttatgtggtccctgtgtgt 4428
Db 4326 AAGAGGAATTTACAAGAGTACTGCAGACATCATGAAATCACTATGTGTGGCCCTGTCTCG 4385
Qy 4429 ataaagaagaccatgtcaagtttaagttcttctcagagaagaagcagacagagaagc 4488
Db 4386 ATAAAGGAAGACCCATGTCAAGGTTTAAGTCTTTTCAGAAAGGAAGGCGACAGAGAAGC 4445

Qy 4489 gtgtgtggagactgaaactgtgtgaccatgtactgcagaaactgaggaactaaagtcactg 4548
Db 4446 GTGTGCTGGAGACTGAACCTGTGTGGACCATGTACTTGCAGAACTGAGGACTAAAGTCACTG 4505
Qy 4549 atgaagaatggcagagaagcttccgataatcttgcagtcgcaaaatctgaaggggtcat 4608
Db 4506 ATGAAGGAATGGCAGAGAAGCTTCCGATAATCTTGCAGTGCAGAAATCTGAAGGGGTCTAT 4565
Qy 4609 ttcttaagtctcaggtttgtttgaaatccatgagcaacagtggttcccatgtgagtg 4668
Db 4566 TTTCTAATGCTTCAGGTTTGTGTAATCCATGAGCAACAGTGGTTCCTCATTTGTGAGTG 4625
Qy 4669 tgcagccccgggagaagctgtcagccagcactagagcgctatgaaactcaggtacaaa 4728
Db 4626 TGCTAGCCCCGGAGAGCTGTACGCCAGCTAGGAGCGCTATGAACTCAGGTACAAA 4685
Qy 4729 ctgacttcagacctcccttgcgaacttaacatcagaagaagcagtgaaattgaaattctgg 4788
Db 4686 CTCGACTTCAGACCTCCCTTCCCACTTACATCAGAAAGCAGTGAATTTCAATTTCTGG 4745
Qy 4789 ctgtgagctaccacaagaacaaatattacagtttttattcattagatggatgctgatg 4848
Db 4746 CTGTGGATCTACCCAAAGAAACAATATTACAGTTTATCATTAGAGTGGATGCTGATG 4805
Qy 4849 aacaggcaatttaacacacactgtgaagcagctgtgtcctgcctgcctgcctgcctgcctgcct 4908
Db 4806 AACAGCATTTAACACAACTGTGAAGCAGCTGTGTGCACGCTGCCCAAGCAAGATACC 4865
Qy 4909 tcaatttagctgtgatgaatttataacatcaaatgtagaagaagaaggtgtctgtgtat 4968
Db 4866 TCAATTTAGTGTGTGATGAAATTTATACATCAAAAGTAGAAAAAAGGTGTGTGTGTAT 4925
Qy 4969 ttctgacagctatagatgactactacagaattcttatttaacccctaaagaactctg 5028
Db 4926 TTCTGTACAGCTATAGAGATGACTACTACAGAACTTATTTTAAACCTTAAAGAACTGTG 4985
Qy 5029 ttaacctattcaaacagacagaggtcttactggaataatggaatgtgtacattcatc 5088
Db 4986 TTAACCTCATTTCAACAGACAGAGGCTTATACTGGAATTAATGGAATCTGTACATTATC 5045
Qy 5089 ataatttaaaatttaattctcagaagaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5148
Db 5046 ATAATTTAAATTTAAATTTCTAAGAGAGGCTGGGTGAGTGGCTCACACCTTTAATCCCA 5105
Qy 5149 gcacttgggagcaagcaggaagctgtgtgaaccagaggtgtgagaccagct 5206
Db 5106 GCATTTGGGAAGCCAGGACGAGGAGTGTCTTGAACACGAGGTTTGAGACCAGCT 5163

RESULT 2

AB037759
LOCUS AB037759 4994 bp mRNA
DEFINITION Homo sapiens mRNA for KIAA1338 protein, partial cds.
ACCESSION AB037759
VERSION AB037759.1 GI:7243056
KEYWORDS
SOURCE

Homo sapiens brain cDNA to mRNA, clone_fhl6948.
Homo sapiens

REFERENCE
1 (sites)

1 Nagase, T., Kikuno, R., Ishikawa, K.I., Hirose, M. and Ohara, O.
Prediction of the coding sequences of unidentified human genes.
XVI. The complete sequences of 150 new cDNA clones from brain which
code for large proteins in vitro

DNA Res. 7 (1), 65-73 (2000)
2 (bases 1 to 4994)

2 Ohara, O., Nagase, T. and Kikuno, R.
Direct Submission

JOURNAL
TITLE Submitted (31-JAN-2000) to the DDBJ/EMBL/GenBank databases. Osamu

Db 3606 TCCCCAGTTTAAAGAGGCGCACAAAGCTCTGGGCGCAGTTCACCTGCCATTTGGGTCAGCAT 3665
Qy 4190 agctataaagaagatatctgctgctgctcctcaacatgaggaatctgttacaataagctc 4249
Db 3666 ACCTATAGACAAAGATATCTGCTGCTGCTCCAACTGAGGAATCTGTTACAATAAGCTC 3725
Qy 4250 ttgtgacctctggttgaagtgtgtgctcagatgtctatgtccagggcccatcaacctaac 4309
Db 3726 TTGTGACCTCTGTTGTAAGTGTGTGTCAGATGCTATGTCTCAGGGCCCATCAACTAAC 3785
Qy 4310 ccagaactctggacacagagcgcacacagcagaaatcatgtacagctggtcagctcca 4369
Db 3786 CCAGAACTCTGGACACAGGATCATCAGCAGAAATCATGTACGACTGTGTACAGTCCA 3845
Qy 4370 agaggaattacaagagctactgcagacatcatgaaatcactatgtacagctggtcagctcca 4429
Db 3846 AGAGGAATTACAGAGTACTGACAGACATCATGAAATCATCTATGTGGCCCTTGTCTCGA 3905
Qy 4430 taaagaggagccatgtcaaggttaagtcttctcgagaaggaaagcagacagagaagcg 4489
Db 3906 TAAAGAAAGGAGGCATGTCAAGGTTAAGTCTTCGAGAAAGGAAAGCAGACAGAGAAGCG 3965
Qy 4490 tgtctgagagactgaactgtgacacatgtactgcagaaactgaggaactaaagtcaactga 4549
Db 3966 TGTGCTGGAGACTGAACCTTGTGGACCATGTACTGCAGAAACTGAGGACTAAAGTCACTGA 4025
Qy 4550 tgaaggaattgagcagaagctcccgataatcttcagtgcaaaatctgaagggtcatt 4609
Db 4026 TGAAGGAATTGCGAGAGAGCTTCGGATAATCTTCGAGTGCAGAAATCTGAAGGGGTCAAT 4085
Qy 4610 ttctaattgctcaggttgttgaatccatgaggaacagtggtccattgtgagtg 4669
Db 4086 TTCTAATGCTTTCAGGTTTGTGTAATCCATCGAGCAACAGTGGTTCCTCCATTTGAGTGT 4145
Qy 4670 gctagcccgagaaagctgctcagcagcactagagggcgtatgaaactcaggttacaac 4729
Db 4146 GCTAGCCCGGAGAGCTGTCAGCCAGACACTAGGAGGCGCTATGAACTCAGGTACAAAC 4205
Qy 4730 tgaactcagacctccctgccaactcatcatcagaaaagcagtgaaattgaaattctggc 4789
Db 4206 TCGACTTCAGACTCCCTTGGCAACTTACATCAGAAAGCAGTGAATTTGAATCTGGC 4265
Qy 4790 tgtgagctaccacaagaacaataattacagtttttattcattagagtggtgagtgatga 4849
Db 4266 TGTGAGTCTACCCAAAGAAACAAATATTACAGTTTTTATCATTTAGAGTGGGATGATGA 4325
Qy 4850 acaggcatttaacacaaactgtgaagcagctgctgcacgctgcacaaagcaagataacct 4909
Db 4326 ACAGGCATTTAACAACTGTGAAGCAGCTGCTGTACGCTGCCAAGCAAGATACCT 4385
Qy 4910 caaattagctgtgatgaattatacatcaaatgtagaagaaagtgctgtgctatt 4969
Db 4386 CAATTTAGTCTGTGATGAATTTTAACTCAATCAAAAGTGAAGAAAGGTGCTGTGCTATT 4445
Qy 4970 tctgtacagctataagatgactactcagaactcttattttaacccttaagaagactgct 5029
Db 4446 TCTGTACAGCTATAGAGATGACTACTACAGAACTTTATTTTAAACCTTAAAGAAAGTCTGCT 4505
Qy 5030 taacctattcaaacagacagaggtcttactggaataatggaatgtgtacattcatca 5089
Db 4506 TAACTCATTTCAACACAGACAGAGCTTATCTGGAATTAATGGAATGTGTACATTCATCA 4565
Qy 5090 taatttaaatattcttaagaagagctgggtgagtggtgagtggtcgaaccttttaaccag 5149
Db 4566 TAAATTAAATTAATTTCTAAGAAAGAGGCTGGGTGAGTGGCTCACACCTTTAATCCAG 4625
Qy 5150 cactttgggaagcgaagcagagactgcttgaacacagaggtttgagacagcctgaag 5209
Db 4626 CACTTTGGGAAGCCAGGAGAGACTGCTGTAACACAGGATTTGACACAGCCTGAG 4685
Qy 5210 caacaagaagaccccatctctataaaaaactaaaaattagttgggcatggtggcaca 5269

Db 4686 CAACAAAGCAAGCCCATCTCTATAAAACTAAAAAATTAGTTGGCATGCTGCACA 4745
Qy 5270 tgcctgtagtccagctactccagaggtgagatgagatcatctgagcctcagaggttga 5329
Db 4746 TGCTGTAGTCCCGAGCTACTCCAGAGGCTGAGATGATCATCTGAGCCTCAGAGGTTGA 4805
Qy 5330 ggtcagctgagctgtgactgcgacctgcactccagctctgggacaacagacaagacc 5389
Db 4806 GGCTGCAGTGAAGTGTGACTGCGCCACTGCACCTCAGTCTGGACACAGACAGACCC 4865
Qy 5390 tgccttaaaaaaagaaaaaatttttttcttaagaagctgctcctcaaaagtga 5449
Db 4866 TGCTTAAAAAAGAAAAAATTTTTTTCTTAAAGAGCTGCTCCTCAAAAGTTGA 4925
Qy 5450 gctttagttagtctcatgtatataataataataataataataataataataataatg 5509
Db 4926 GCTTTGTTAGTTTTCATGTGTATATATATATATATATATATATATATATATATAT 4985
Qy 5510 ctttcatat 5518
Db 4986 CTTTCATAT 4994

RESULT 3

LOCUS HSM802494 3351 bp mRNA PRI 18-FEB-2000
DEFINITION Homo sapiens mRNA; cDNA DKFZp434F1312 (from clone DKFZp434F1312);
partial cds.
ACCESSION AL157497
VERSION AL157497.1 GI:7018544

KEYWORDS human.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 3351)

AUTHORS

Duesterhoeft A., Lauber J., Mewes H.W., Weill B. and Wiemann S.

TITLE

Direct Submission

JOURNAL

Submitted (18-FEB-2000) MIPS, Am Klopferspitz 18a, D-82152

COMMENT

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the CDNA sequencing
consortium of the German Genome Project.
This clone (DKFZp434F1312) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at <http://www.mips.biochem.mpg.de/proj/cdna/>.

FEATURES

Location/Qualifiers
1..3351
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp434F1312"
/DH10B; sites NotI + SalI"
/dev_stage="adult"
/tissue_type="testis"
1..2819
/gene="DKFZp434F1312"
1..2819
/note="similarity to eIF-2alpha kinase (Drosophila
melanogaster)"
/codon_start=3
/product="hypothetical protein"
/protein_id="CAB75678.1"
/db_xref="GI:7018545"
/translation="STSGERSASARFATPGSSDDEDDDEHGVFSQSFLPASDS
ESDIIFDNEDENSKONQDCNEKNGCHESEPSVTTEAVHYLIQMEYCEKSTLRDT
IDQGLYRDTVRLWRLFRLEILDGLAYIHKGMIHRLKPNVIFLSDSDHRLIGDGLAT
DLHAFSADSKDDQDTGLKSDPSGHLTGMVGTALYVPEVGSTKSAYNQKVDLFL
GLIFFEMSHPMVTASERIFVLNQLRDTSPKFPEDFDGGEHAKQKSVISWLLNHDA

gene

CDS

KRPTATELLKSELLPPPMQMESESELHEVLHHTLTNVWDGKAYTMTMAQIIFSORISPAIDI
 TDSIDLGNFSIRITAKMCHVCETIIRFKHGAVOLCTPLLPKRNQIYEHNEARAL
 FMDHSGMLKMPEDLPAPRYVARNNLNKRKYCIERVPKLDHDPKHEKCECAF
 DIVSTYNSFTPEAKIITYETIOEPFALQBNRYSYNLNHTMLKALLHCHGIPEDK
 LSOVYITIDAVTAEKTYTREVREKCFNCSLNUCLYKTEIQGLQDLMTPEIDINL
 TKQGTGAOLKYKGLDEVLVGLKGLQVLPGLVINGVYQVQINGIYFQVAPRI
 KRQRAPVEILAGSGRYDILLIPQFRGPAQLGAPYICGSTAIDKISAALVINESVYT
 ISSCDILTVSGVQMSRNDLNTKULMTAGITAEITMYDMSQSOEDELQYCFRHEITLV
 ALVSDGEGSHVYKSEKEHQETREKRVLELTDHVLHQLKRYTDOERNGREASDNTAV
 QNLKQFSNAGSLFEIHTAGTVYVSLVAPELKSLSRRRTYETQVTRQIQTISLNLHQ
 KSESEILAILDPLFKYSRDIYDADQAPNTYTKOLLKSLPKQRYLYKULVCDEITYN
 TKVEKYSVLFLKSYRDIYDYL*"

polvA signal

polyA_signal
polyA_site

polya_size	3334	
BASE COUNT	1017 a	734 c
	764 q	836 t

[illegible]

Query Match	60.3%;	Score 3332;	DB 93;	Length 3351;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3332;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

Qy	2194	ggagcacttggcgagcgctggcgagctgcccgtttcccoogcaccgcccggggtcca	2253
Db	1	GGAGCATTCTGGCGGAGCGCTGGCGAGTGGCCGCTTTCCCGCACCGCCGGGTCCA	60
Qy	2254	gcgatgcagaggaacgcagcaagagcagcagctggcgcttctccagctcctctgc	2313
Db	61	CGCATGACGAGGACGACGAGGAGACGACCGTGGCGCTTCTCCAGTCTCTCTGC	120
Qy	2314	ctgcttcagattctgaagtgatattatctttgacaatgaagatgagacagataaagtc	2373
Db	121	CTGCTTCAGATTCTGAAAGTGATATTATCTTTGACAATGAAGATGAGAACAGTAAAGTC	180
Qy	2374	agaatcaagatgaagattcaatgaaagaaagctgctgccatgaagtgagccatcagtga	2433
Db	181	AGAATCAGGATGAAGATTGCATGAAAGAAAGATGGCTGCCATGAAGATGACCATCAGTGA	240
Qy	2434	cgaatgggctgtgcactactatcacatccagatggagtctgtgagaagagcacttac	2493
Db	241	CGACTGAGGCTGTGCACACTCTATACATCCAGATGGAGTACTGTGAGAAGAGCACTTTAC	300
Qy	2494	gagacacattgaccagggaactatcagagacacgcgtcagactctggagggcttttcgag	2553
Db	301	GAGACACATTGACGAGGACTGTATCGAGACACCGTCAGACTCTGGAGGCTTTTTCGAG	360
Qy	2554	agattctggatggattagcttatccatgagaaagggaatgattcacccgggatttgaagc	2613
Db	361	AGATTCTGGATGGATTAGCTTATATCATGAGAAAGAAATGATTCACCGGATTTGAAGC	420
Qy	2614	ctgtcaacattttttggattctgatgaccatgaaataggatgatttggtttggcga	2673
Db	421	CTGTCAACATTTTTTTGGATCTGTATGACCACTGTAATAGTGTGATTTGGTTGGCGA	480
Qy	2674	cagaccatctagccttctctgtcagacaacaacagcagatcacagagagacttgatta	2733
Db	481	CAGACCATCTAGCCTTTTCTGCTGACAGCAAAACAGCAGATCAGACAGAGACTTGATTA	540
Qy	2734	agtcagacccttcaggctcaacttaactgagatgttgccactctctctatgtaagccag	2793
Db	541	AGTCAGACCCCTTCAGGTACCTTAATCGGATGGTGTGGCACTGCTCTATGTAAAGCCAG	600
Qy	2794	aggtccaaggaagcaccacaaatctgtcatcacaccagaagtgatctcttcagcctggaa	2853
Db	601	AGGTCCAAGGAAGCACCAAAATCTGCATACACCAAGATGGATCTCTTCAGCCCTGGAA	660
Qy	2854	ttatcttttgagatgtcctatcacccatgttcagcgttcagaagatctttgttc	2913
Db	661	TTATCTCTCTTTGAGATGTCCTATCACCCCATGGTTCAGGCTTCAGAAAGATCTTTGTTC	720
Qy	2914	tcaaccaactcagatcccacttcgcctaagtttcagaagaactttgacgatggagagc	2973
Db	721	TCAACCTCACTCAGATCTCCACTTCGCTTAAGTTTCAGAAAGACTTTGACGATGGAGAGC	780

		BASE COUNT	921 a	641 c	663 g	761 t	
		ORIGIN					
		Query Match	48.0%;	Score 2650;	DB 89;	Length 2986;	
		Best Local Similarity	99.9%;	Pred. No. 0;			
		Matches 2850;	Conservative	0;	Mismatches	4;	Indels 0; Gaps 0;
2y	2553	gagattctgagattgagcttatccatcagagaagggaatgattccacgggattggaag	2612				
Db	1	GAGATTCTGGATGGATTAGCTTATATCCATGAGAAAGGAATGATTCACCGGGATTGAAG	60				
2y	2613	cctgtcaacatttttttgattctgatgaccatgtgaaatagggtgatttggttgagc	2672				
Db	61	CTGTCAACATTTTTTTGGATTTTGATGACCATGTGAAATAGGTGATTTTGGTTGGCG	120				
2y	2673	acagaccatctagcttttctgtcagacaaacagacgattcagacgagagacttgatt	2732				
Db	121	ACAGACCATCTAGCGCTTTCTTGCTGACAGCAACACAGACGATCAGACAGAGACTTGATT	180				
2y	2733	aagtcagacccttcaggtcacttaacttggaatgggttgggaactgctctctatgtaagcca	2792				
Db	181	NAAGTCAGACCCCTTCAGGTTCATTAACTGGGATGGTTGGCACTGCTCTCTATGTAAGCCCA	240				
2y	2793	gaggtccaagaagacacacacatctgcatacaaccagaaagtggatctcttcagctggga	2852				
Db	241	GAGGTCCAAGGAGACCCAAATCTGCATACCAACAGAAATGGATCTCTTCAGCCCTGGGA	300				
2y	2853	attatctcttttgagatgtcctatcaaccatgggtcagcgcttcagaagaagactcttgt	2912				
Db	301	ATTATCTCTTTTGAGATGCTCTATCACCCATGCTACGGCTTCAGAAAGATCTTTGTT	360				
2y	2913	ctcaaccaactcagagatcccaactcgcttaagtttccagaagaacttgcagatggagag	2972				
Db	361	CTCAACCAACTCAGAGATCCCACTTCGCGCTTAAGTTTCCAGAAAGACTTTGACGATGGAGAG	420				
2y	2973	catgcaaaagcaaatcagtcactctctgtgctgttgaacacgactcagcaaacgccc	3032				
Db	421	CATGCAAAAGCAAAATCAAGTCATCTCTGGGTGTTTGAACCAACGATTCAGCAAAACGCCCC	480				
2y	3033	acagccacagaactgctcaagsgtgagctgctgtcccccaacccagatggagagagtcagag	3092				
Db	481	ACAGCCACAGAACTGCTCAAGAGTGAGCTGTGCTGCCCCACCCACAGATGGAGAGTCAAG	540				
2y	3093	ctgcatgaagtgctgcacacacgctgaccaaactgtgagtggaagagcctaccgaccatg	3152				
Db	541	CTGCATGAAGTGTGCACACACGCTGACCAACGTGGATGGGAAGGCCCTACCGCACCATG	600				
2y	3153	atgcccagatctcttcgacgcgcatctccctcgcatcgattcacacctatgacagcag	3212				
Db	601	ATGCCCCAGATCTCTTCGACGCGCATCTCCCTCGCATCTGATTACACCTATGACAGCGAC	660				
2y	3213	atactgaagggcaactctctcaactcgtacagccaagatgcagcacatgtgtgtgaaacc	3272				

QY 4353 gactgtcacagtcccaagagaataacaaagtactcagacatcatgaaatcacctat 4412
 DB 1801 GACTGGTACAGTCCCAAGAGAAATTACAGAGTACTCAGACATCATGAATCACCTAT 1860
 QY 4413 gtggccctgtctcgataaagaagaagccatgtcaaggttaagtcttctcgaaagaa 4472
 DB 1861 GTGGCCCTGTCTCGGATAAAGAAGGAGCCATGTCAAGGTTAAGTCTTTCGAGAGGAA 1920
 QY 4473 aggcagacagagaacggtgtgctggagagctgaactgtgagccatgtactgcagaaactg 4532
 DB 1921 AGGCAGACAGAGAAGCGTGTGCGAGACTGAACCTGTGGACCATGTACTGCAGAAACTG 1980
 QY 4533 agdactaaagtcactgatgaagaagatggcagagaagcttcccgataatcttgcagtcac 4592
 DB 1981 AGGACTAAGTCACTGATGAAGGATGGCAGAGAGCTTCCGATATATCTTGCAGTGCAA 2040
 QY 4593 aatctgaagggttcattcttaagtcttcaggttggttgaaatccatggagcaacagt 4652
 DB 2041 AATCTGAAGGGGTCAATTTCTAATGCTTCAGGTTGTGTAATCCATGGAGCAACAGTG 2100
 QY 4653 gtcccatgtgagtgctagcccgagagagctgtcagccagcactagagcgctat 4712
 DB 2101 GTTCCATGTGAGTGTGCTAGCCCGGAGAGCTGTACGACGACTAGGAGCGCTAT 2160
 QY 4713 gaaactcaggtacaaactcgacttcagaccccttgcacactacatcagaaagcagt 4772
 DB 2161 GAAACTCAGGTACAACTCGACTTCAGACCTCCCTGCCACATTCATCATCAAGAGCAGT 2220
 QY 4773 gaaattgaaattctggctgtgagctacccaaagaaacaaatattacagtttttcatca 4832
 DB 2221 GAAATTGAAATTCTGGCTGTGGTATCTACCCAAAGAAACAATATTACAGTTTTATCATTA 2280
 QY 4833 gagtggagtgctgatgaacaggtattacaacactgtgaagcgtgtgcacgctg 4892
 DB 2281 GAGTGGATGCTGATGACAGGCAATTAACACAACCTGTGAAGCAGCTGTGTACGCGCTG 2340
 QY 4893 ccaagcagaagatcactcctaaattagctgtgtgataaattataacatcaaaagttagaaaa 4952
 DB 2341 CCAAGCAAGAGATACCTCAAAATTAGTGTGATGAATTTAATCAATCAAGTAGAAAAA 2400
 QY 4953 aagtgctgtgctattctgtacagctatagagtagactactacagaactctatttaa 5012
 DB 2401 AAGTGTCTGTGCTATTCTGTACAGCTATAGAGTACTACTACAGAACTTTATTTTAA 2460
 QY 5013 ccctaaagaagctgcgttaacctattcaacacagacagagcgttatactagaataatgga 5072
 DB 2461 CCCTGAAGAACTGTCATTAACTCATTCAACACAGACAGAGCTTATCTGGAATATATGA 2520
 QY 5073 atgtgtacattcatataatttaaaatttaaaattcttaagaagaggtgggtgcagtggt 5132
 DB 2521 ATGTGTACATTTCATATAATTTAAATTTAAATTTCTAAGAAAGAGGCTGGGTGCAGTGGCT 2580
 QY 5133 cacaccttaatccagcactttggaagccaaagcaggaagactgcttgaaccaggag 5192
 DB 2581 CACACCTTTAATCCACACTTTGGGAAGCCAGGAGGAAAGACTGCTGAAACCAAGGAG 2640
 QY 5193 ttgagaccagcctgagcaacaaagcaagaccctctctataaaactataaaataatag 5252
 DB 2641 TTTGAGACCAAGCCTGAGCAACAAGCAAGACACCCCATCTCTATAAAACTAAAAAATAG 2700
 QY 5253 ttgggaatgtgcacatgctgtagtcaccagctactccagagaggtgagatgatcatct 5312
 DB 2701 TTGGGCATGTGTGCATATGCTGTAGTCCAGCTACTCCAGAGGCTGAGATGATCATCT 2760
 QY 5313 gagcctcagaggttgaggtgcagtgagctgtgactgcagcactgcactccagctggg 5372
 DB 2761 GAGCCTCAGAGGTTGAGGCTGCAGTGAGCTGTGACTTGCCCACTGCATCCAGTCTGGG 2820
 QY 5373 acaacagagcaagaccctgtcttaaaaaa 5406
 DB 2821 ACAACAGAGCAAGACCCTGTCTTTAAAAA 2854

RESULT 5

HSA243428

LOCUS

DEFINITION

Homo sapiens partial mRNA for putative eIF2 alpha kinase (GCN2

gene)

AJ243428

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

polya_signal

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 2161; Conservative

QY 3364 ctgccttattatgacacagcgagctgtggtgatgcttcttgcacctggatcc 3423

DB 1 CTGCCCTATTATGATGACACAGCGGATGCTGCTGCTCTCTTTTACCTGCGGATCC 60

QY 3424 ctttgcagatatgtgcaagaaataatattgaatttaaaacagactgcatgataac 3483

DB 61 CTTTTCGAAGATATGTCGCAAGAAATATATTATTATTTAAACGATACCTGATAGAAC 120

QY 3484 gtgtgtcagcgccgcaagttagatcgatttcacccaaagaaactcttgagtgatgcat 3543

DB 121 GTGTGTTTCAGGCGCGCAAGTTAGATCGATTTCATCCCAAGAACTTCTGGAGTGTGCAT 180

QY 3544 ttgatattgtcacttctaccaccacagcttctgcccactgctgaaattatctacacta 3603

DB 181 TTGATATTGTCACTTCTACCAACACAGCTTTCTGCCACTGCTGAAATTTATCTACACTA 240

HSA243428 2162 bp mRNA PRI 15-OCT-1999
 Homo sapiens partial mRNA for putative eIF2 alpha kinase (GCN2 gene)

AJ243428
 AJ243428.1 GI:6065913

eIF2 alpha kinase; GCN2 gene.
 human.
 Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2162)

Berlanga, J.J., Santoyo, J. and De Haro, C.
 Characterization of a mammalian homolog of the GCN2 eukaryotic initiation factor 2alpha kinase
 Eur. J. Biochem. 265 (2), 754-762 (1999)
 99435990
 2 (bases 1 to 2162)

Santoyo, J.
 Direct Submission
 Submitted (28-JUN-1999) Santoyo J., Molecular Biology, Centro de Biologia Molecular 'Severo Ochoa', Fac. Ciencias. U.A.M., Cantoblanco, Madrid, 28049, SPAIN
 Location/Qualifiers

1..2162
 /organism="Homo sapiens"

/db_xref="taxon:9606"

1..2144
 /gene="GCN2"

<1..1649
 /gene="GCN2"

/function="protein synthesis control"

/codon_start=3
 /product="putative eIF2 alpha kinase"

/protein_id="CAB58360.1"

/db_xref="GI:6065914"

/translation="ALFMDHSGMLVLPDFLRIFARVAVARNILILKRYCIERVFRP
 RKLDRFPELLECAFDIVTSTNSFLPTAEIITVIEIIOFFPALGRNYSIYLNT
 MLKAILLHCIGIPEDKLSQVYIILYDAVTEKLRREVAKFCNLSLSNLSLYKFI
 EOKGDLQDMPTINSLIKQGTQIAQLVKYGLKLEEVVGLKLGKLVGLV
 KVOHNGIIFQVAFIKRRRAVPEILAAAGGYDLLIPQFRGALGPVPTAIGVSTA
 IDKLSAVALNMEESVTISSCDLLVSVGMSMSRAINLTQKLTAGTAEIWDWSQS
 QEELEQYCRHHEITVYALVSKESHVKVSKFERQTERKRVLELVDHVLQKLRK
 VIDERNGREASONLAVQNLKGSFNASGLFHGATVVIIVSLAPEKLSKRRRYE
 TVOTRIQTSLANLHQKSETEILAVDLPKETILQFLSLEWDADAQAFNTVKKLLSR
 LPKORYLKLVDCEIYINIKVKKSVLYFLYSYRDDYRILF"

polya_signal

BASE COUNT 672 a 434 c 481 g 575 t

ORIGIN

Query Match 38.2%; Score 2111; DB 93; Length 2162;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2161; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3604 tctatgaataatcaccagaagttccagcacttcagaaagaataattacagtattatttga 3663
 Db TCTATGAATCATCCAGAGATTTCCAGCACTTCAGGAAAGAAATACAGTATTAFTTGA 300
 QY 3664 accataccatgttatgaagcaatactctacactgtggatcccagaagataaactca 3723
 Db ACCATACCATGTTATTGAAGCAATCTCTTACACTGTGGGATCCCAAGAGATAACTCA 360
 QY 3724 gtcaagtctacattattctgtatgtctgtgacagagaagctgacagagagaagtg 3783
 Db GTCAAGTCTACATTTATTTCTGTATGCTGTGACAGAGAAGCTGACGAGGAGAGAAGTGG 420
 QY 3784 aagctaattttgaactgtcttcttcttaataagtcgtgtcgaactcaagattta 3843
 Db AAGCTAAATTTGTAATCTGTCTTTGCTTCTAATAGTCTGTCTGACTCTCAAGATTTA 480
 QY 3844 ttgaacagaaggagatttgaagatttgcacacattatgcacaaataaaattcatttaataaacaga 3903
 Db TTGAACAGAGAGGAGATTTCAGAGATCTTATGCCARCAATTAATTCATTAATAAACAGA 540
 QY 3904 aaacaggtattgcacagttggtgaagtatggcttaaaagacctagagaggttcttggac 3963
 Db AAACAGGTATTGCACGTGGTGAAGTATGGCTTAAAGACCTTAGAGGAGGTTGTTGGAC 600
 QY 3964 tgttgaagaactcggcatcaagttacagctcttgatcaatttggcttggcttacaagg 4023
 Db TGTGGAAGAACTCGGCATCAAGTTACAGGTCTTGATCAATTTGGGCTTGGTTTACAGG 660
 QY 4024 tgcagcagcaaatggaaatcctctccagtttggcttttcatcaaacgaagcaagg 4083
 Db TGCAGCAGCACAAATGGAAATCATCTCCAGTTTGGCTTTCATCAACGAAGCAAGGG 720
 QY 4084 ctgtacctgaatcctcgcagctggagcagatgatgacctgctgattccccagtttag 4143
 Db CTGTACCTGAAATCCTCGCAGCTGGAGGAGATATGACCTGCTGATTCGCCAGTTTAG 780
 QY 4144 ggcacaagctcgtggccagttccactgccattgggtcagcatagctatagacaaga 4203
 Db GGCCAAAGCTCTGGGGCCAGTCCCACTGCCATGGGGTCAGCATAGCTATAGACAAGA 840
 QY 4204 tatctgctgtcgtcctcaacatgaggaatctgttacaataagctcttgaacctctgg 4263
 Db TATCTGCTGCTGCTCAACATGAGGAAATCTGTACAATAAGCTCTGTGACCTCTCGG 900
 QY 4264 ttgtaagttgttgtagatgtctatgtccaggccatacaacctaaaccagaactctgga 4323
 Db TTGTAAAGTGTGGTCAGATGCTATGTCCAGGGCCATCAACCTAACCCAGAACTCTGGA 960
 QY 4324 cagcagcatcacagcagaataatcatgtacactgtgtcacagtcccaagagaaattacaag 4383
 Db CAGCAGGATCATACAGCAAAATCATGTACGACTGTCACAGTCCCAAGAGGAATTACAAG 1020
 QY 4384 agtactgcagacatcatgaaatcacctatgtggcccttgtctcggataaagaagaagcc 4443
 Db AGTACTGCAGACATCATGAAATCACCTATGTGGCCCTTGTCTCGGATAAAGAGGAAGCC 1080
 QY 4444 atgtcaagtttaagcttctcagaagaagaaagcagacagaagcgtgtcgtgagactg 4503
 Db ATGTCAAGGTTAAGTCTTTGAGAAAGAAAGGCACAGAGAGCGTGTCTGGAGACTG 1140
 QY 4504 aacttgtgaccatgtactgcagaactgagactaaagtacactgatgaaaggaattgga 4563
 Db AACTTTGGACCATGTACTGCAGAAACTGAGGACTTAAAGTCACTGATGAAAGGAATGGCA 1200
 QY 4564 gagaagttccogataatcttgcagtgcaaaatctgaaggggtcattttcattatcttcag 4623
 Db GAGAAGCTTCCGATATCTTGCAGTGCAAAATCTGAAGGGGTCTATTTCTAATGCTTCAG 1260
 QY 4624 gtttcttgaatccatgagcaacagtggtcccatctgtgagtgctagccccggaga 4683
 Db GTTTGTTGAAATCCATGGAGCAACAGTGGTTCCTCATTTGTGAGTGTGTAGTCCCGGAGA 1320
 Db 1261 GTTTGTTGAAATCCATGGAGCAACAGTGGTTCCTCATTTGTGAGTGTGTAGTCCCGGAGA 1320

QY 4684 agctgtcagcagcactagaggcgctatgaaactcagggtacaaactcagacttcagacct 4743
 Db ACTGTACGCCAGCACACTAGAGGCGCTATGAAACTCAGGTACAAACTCGACTTCAGACCT 1380
 QY 4744 cccctgcacaattacatcagaaaaagcagtgaaattgaaattcctgctgtgagctaccca 4803
 Db COCTTGCCCAACTTACATCAGAAAAAGCAGTGAATTTGAATTTCTGGCTGTGGATCTACCCA 1440
 QY 4804 aagaacaataattacagtttttattcattagagtggtgagtgagaaagcagcatttaaca 4863
 Db AAGAAACAATATTACAGTTTTTATCATTTAGAGTGGATGCTGATGAACAGAGCATTTAACA 1500
 QY 4864 caactgtgaagcagctgtctcagcctgcagcctgcagcctgcagcctgcagcctgcagcct 4923
 Db CAACCTGTGAAGCAGCTGTCTGTCACGCTGCCAAAGCAAGATACCTCAAAATTAGTCTGTG 1560
 QY 4924 atgaattattataacatacaagtagaaaaaagggtgtctgtctatttctgtacagctata 4983
 Db ATGAATTTATTAACATCAAGTAGAAAAAGGTGCTGTGCTATTCTGTACAGCTATA 1620
 QY 4984 gagatgactactacagaatctttatttaacccctaaagactgtgttaacctcattcaaa 5043
 Db GAGATGACTACTACAGAACTCTATTTTAACCTTAAGAACTGTCTGTTAACCTCATTCAAA 1680
 QY 5044 cagacagaggttatacttggaataatggaaatgtgtacattcattcaataatttaaaattaa 5103
 Db CAGACAGAGGCTTATACCTGAATAATGGAATGTTGATATTTCATCATTAATTTAAATTTAA 1740
 QY 5104 attctaagaagaggtgggtgcagtggtgcacacctttaaaccagcacttgggaagcc 5163
 Db ATTCTAAGAAGAGGCTGGGTGAGTGGCTCACACCTTTAATCCCAAGCACTTTGGGAAGCC 1800
 QY 5164 aaggcaggaagactgcttgaaccagaggtttgagaccagcctgagcagcaacaagaagac 5223
 Db AAGGCAGGAAGACTGCTTGAACACAGGAGTTTGAGACCAGCTGAGCAACAAGCAAGAC 1860
 QY 5224 cccatctctataaaaaactaaaaaattagttggcagcctgagcagcctgagcagcagcagc 5283
 Db CCCATCTCTATAAAAACTAAAAAATTTAGTTGGCATGGTGGCATGCTGTAGTCCCA 1920
 QY 5284 gctactcagagctgagatgatcatctgagcctcagaggttgagcctgagcagcagcagc 5343
 Db GCTACTCCAGAGCTGAGATGATCATCTGAGCCTCAGAGGTTGAGGCTGCAGTGAAGCT 1980
 QY 5344 gtgactgcgcccactgcactccagctgggacacagagcaagcagcagcagcagcagcagc 5403
 Db GTGACTGCGCCACTGCACCTCCAGTCTGGGACAAACAGAGCAAGACCCCTGTCTTAAAAAAA 2040
 QY 5404 aaagaaaaaatttttttttctagaagcgtcctcacaagttgagcctttagtttt 5463
 Db AAAGAAAAAATTTTCTTAAGAAAGCTGTCCCTACAAAGTTGAGCTTTGTTAGTTTT 2100
 QY 5464 tcatgttaataattataaatttctttgggataataataaagccttctcattacacgtg 5523
 Db TCATGTATATATTATTAATTAATTTATCTTTTGGGATATAATAATGCTTTTCATATACCTG 2160
 QY 5524 ca 5525
 Db CA 2162

RESULT 6

HSM802391

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

HSM802391 1138 bp mRNA

Homo sapiens mRNA; cDNA DKFZp434P0612 (from clone DKFZp434P0612);

partial cds.

AL137627

AL137627.1 GI:6808424

human.

Homo sapiens

Eukaryota;

Metazoa;

Chordata;

Craniata;

Vertebrata;

Euteleostomi;

Mammalia;

Eutheria;

Primates;

Hominidae; Homo.

18-FEB-2000

REFERENCE 1 (bases 1 to 1138)
 AUTHORS Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
 TITLE Direct Submission
 JOURNAL Submitted (15-JAN-2000) MIPS, Am Klopferspitz 18a, D-82152 Martinsried, GERMANY
 COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by Oligen (Hilden/Germany) within the cDNA sequencing consortium of the German Genome Project.
 This clone (DKFZp434P0612) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://www.mips.biochem.mpg.de/proj/cDNA/.

FEATURES
 Location/Qualifiers
 1..1138
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="DKFZp434P0612"
 /clone_lib="434 (synonym: htes3). Vector pSport1; host DH10B; sites NotI + SalI"
 /dev_stage="adult"
 /tissue_type="testis"
 1..576
 /gene="DKFZp434P0612"
 <1..576
 /note="DKFZp434P0612"
 /note="similarity to C-terminus of Drosophila eIF-2alpha kinase"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CAB70849.1"
 /db_xref="GI:6808425"
 /translation="GSHVKVKSFEKERQTEKRVLETVLVDHVLQKLTKVTDERNGRE ASDNLAVQNLKGSFNSAGLFEIHGATVPIVSLAPEKLSASTRRYETQVOTRLQT SLANLHOKSSEILELAVDLKPEILQFLSLEWDADEQAFNTVKKLLSLRPLKQRYLKL VCEIYNKVKVSVLFLYSRDDYRILF"
 polyA_signal 1066..1071
 polyA_site 1078
 BASE COUNT 410 a 207 c 242 g 279 t
 ORIGIN

Query Match 18.78; Score 1032; DB 93; Length 1138;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1082; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4437 ggaagccatgtcaagggttaagctcttcgaagaagaaaggcagacagagaagcggtgctg 4496
 |||||
 Db 1 GGAAGCCATGTCAAGGTTAAGTCTTTTCGAGAGGAAAGGCGAGACAGAGCGGTGCTG 60
 |||||

Qy 4497 gagactgaactgtggaccatgtactcagaactgaggactaaagtcaactgatgaaagg 4556
 |||||
 Db 61 GAGACTGAACCTGTGGACCATGTACTGCAGAACTGAGGACTAAAGTCACTGATGAAGG 120
 |||||

Qy 4557 aatggcagaagcttcgcataatcttcgagtcgaaactgaaaggggtcatttctaatt 4616
 |||||
 Db 121 AATGGCAGAGAAGCTTCGATAATCTTGCAGTGCAGAAATCTGAAGGGGTCAATTTCTAAT 180
 |||||

Qy 4617 gcttcagggtttgttgaatccatggagcaacagtggttccattgtgagtgctagcc 4676
 |||||
 Db 181 GCTTCAGGTTGTTGAAATCCATGGAGCAACAGTGGTTCCTCCATTGTGAGTGGCTAGCC 240
 |||||

Qy 4677 ccggagaagctgtcagccagcactagagcgctatgaaactcaggtacaaactcgactt 4736
 |||||
 Db 241 CCGGAGAAGCTGTGAGCCAGCCTAGGAGCGCTATGAACCTCAGGTACAAACTCGACATT 300
 |||||

Qy 4737 cagactcccttgcgaacttacatcagaaagcagtgaaattgaaattcttgctggat 4796
 |||||
 Db 301 CAGACCTCCCTTGCACACTTACATCAGAAAGCAGTGAATTTGAAATTTCTGGCTGGAT 360
 |||||

Qy 4797 ctacccaagaacaataattacagttttttatcatgagtgagtgatgatgaacaggca 4856
 |||||

Db 361 CTACCCAAAGAAACAATATTACAGTTTTTATCATATTAGATGGGATGTGATGAACAGCA 420
 |||||

Qy 4857 tttaacaaactgtgaagcagctgtctcagcgcctgccaaagcaaaagataacctcaaat 4916
 |||||

Db 421 TTTAAACAACACTGTGAAGCAGCTGCTGTACAGCTGCCAAGCAAGATACCTCAAAATTA 480
 |||||

Qy 4917 gtctgtgatgaatttataacatacaaaagtgaaaaaagtgctgtctgtattctgtac 4976
 |||||

Db 481 GTCTGTGATGAATTTTATAACATCAAAAGTAGAAAAAGGTGCTGTGCTGCTATTCTGTAC 540
 |||||

Qy 4977 agctatagatgactactacaqaactctattttaaccctaaagaactgcgttaacctc 5036
 |||||

Db 541 AGCTATAGATGATGACTACTACAGAACTCTTATTTTAAACCCCTAAAGAACTGTCGTTAACCTC 600
 |||||

Qy 5037 attcaaacagacagaggtcttatactggaataatggaatgtgtacattcatcataatta 5096
 |||||

Db 601 ATTCAAAACAGACAGAGGCTTATACTGGAATAATGGAATGTTGTACATTTCATCATATTTA 660
 |||||

Qy 5097 aaattaaattctaagaagagctgggtgcagtgagtggtgcacaccttttaaccagcaacttg 5156
 |||||

Db 661 AATTAATTAATTTCTAAGAGAGGCTGGGTGCAGTGGCTCACACCTTTTAAATCCACGACTTTG 720
 |||||

Qy 5157 ggaagccaagcgaggaagactgcttgaaaccagaggtttgagaccagcctgagcaacaaa 5216
 |||||

Db 721 GGAAGCCAAAGCGAGGAGACTGCTTGAACCCAGAGAGTTTGAGACACAGCCTTGACCAACAA 780
 |||||

Qy 5217 gcaagaccccatctctataaaaaactaaaaaatttagttggcagtggtgcacactgctgt 5276
 |||||

Db 781 GCAAGACCCCATCTCTATATAAAAACTAAAAAATTAGTTGGGCATGCTGACACATGCTGT 840
 |||||

Qy 5277 agtccagctactccagagctgagatgagatcctgagcctcaggaggttgaggtgca 5336
 |||||

Db 841 AGTCCAGCTACTTCCAGAGGCTGAGATGATCATCTGAGCCTCAGAGGTTTGAGGGTGCA 900
 |||||

Qy 5337 gtgagctgtgactgcgccactgcactccagctcctgggcaacagagcaagacctgtctta 5396
 |||||

Db 901 GTGAGCTGTGACTGCGCCACTGCACCTCCAGTCTGGGCAACAGAGCAAGACCTGTCTTA 960
 |||||

Qy 5397 aaaaaaaagaaaaaaatttttttctaagaagctgctctacaaaagtgtgagctttgt 5456
 |||||

Db 961 AAAAAAAAGAAAAAATTTTCTTAAGAGAGCTGCTCTACAAAGTGTGAGCTTTGT 1020
 |||||

Qy 5457 tagttttcatgtgtaataattataattttatcttttggatataataaagtcttcat 5516
 |||||

Db 1021 TAGTTTTTCATGTGTAATATATATATAAATTTATCTTTTGGGATATAATAAATGCTTTTAT 1080
 |||||

Qy 5517 ata 5519
 |||

Db 1081 ATA 1083

RESULT 7
 HSM802022 1917 bp mRNA PRI 18-FEB-2000
 LOCUS Homo sapiens mRNA; cDNA DKFZp434H149 (from clone DKFZp434H149).
 DEFINITION AL137676
 ACCESSION AL137676
 VERSION AL137676.1 GI:6807846
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1917)
 AUTHORS Poustka,A., Klein,M., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
 TITLE Direct Submission
 JOURNAL Submitted (15-JAN-2000) MIPS, Am Klopferspitz 18a, D-82152 Martinsried, GERMANY
 COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by DKFZ (German Cancer Research Center, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.

This clone (DKFZp344H149) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de further
information about the clone and the sequencing project is available
at <http://www.mips.biochem.mpg.de/proj/cDNA/>.

FEATURES
Source
1..1917
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp344H149"
/clone_lib="434 (synonym: htes3). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
/tissue_type="testis"
1891..1896
polyA_signal
polyA_site
BASE COUNT 612 a 376 c 416 g 513 t
ORIGIN

Query Match 16.68; Score 919; DB 93; Length 1917;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1019; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3431 aagatatgtgcaagaataataattgaatttaaaacgatactgcatagaacgtgttt 3490
Db 1 AAGATATGTGCAAGAAATATATATATGAATTTAAACGATACGCATAGAACGTGTGTT 60
QY 3491 caggccgcgaagttagatcgatttcacccaaagaactcttgagtgatgattgat 3550
Db 61 CAGGCCGCCGAAGTTAGATCGATTTTCATCCCAAGAACTTCTGGAGTGTGCAATTGTAT 120
QY 3551 tgcactctaccacaaacagcttctccacactgctgaattatctacactatctga 3610
Db 121 TGTCACTTCTACCAACACAGCTTTCTGCCACTGCTGAAATTTATCTACATATCTATGA 180
QY 3611 aatcatccaagagttccagcacttcaggaaagaataatcacagtatttttgaaaccatac 3670
Db 181 ATATCCCAAGAGTTCCAGCATTCCAGAAAGAAATTTACAGTATTATTGTAACCATAC 240
QY 3671 catgtattgaagaacaactcttactacgtggtggtatcccaagaataaactcagtcagt 3730
Db 241 CATGTATTGAAGCAATACCTTTACACTGTGGATCCCAAGATAAATCACTCAAGT 300
QY 3731 ctacattattctgatgtctgtacagaaagctgcagggagagaaagtggagctcaa 3790
Db 301 CTACATTATCTGTATGATGCTGTGACAGAGAGCTGACAGGAGAGAGAGTGGAGCTAA 360
QY 3791 atttgtaactgtcttcttcttcttaatagtctgtgtgactctacaaagtttattgaaca 3850
Db 361 ATTTGTAATCTGCTTTGCTTCTTAATAGTCTGTGCTGACTCTACAAGTTTATTGAACA 420
QY 3851 gaaggagagattgcgaagatcttatcccaacaataaattcattataaaacagaaacagg 3910
Db 421 GAAGGAGAGATTGCAAGATCTTATGCCAACATAAATTCATTAAATAAACAGAAACAGG 480
QY 3911 tattcacagttggtgaagatggtctaaagacactagagaggtgtgtgactgttga 3970
Db 481 TATTGCACAGTTGGTGAAGTATGGCTTAAAGACCTAGAGAGGTTGTGGACTTTGCA 540
QY 3971 gaaactcggcatcaagttacaggtcttatcaatttggtggttggtttacaaggtgcagca 4030
Db 541 GAAACTCTGCATCAAGTTACAGTCTTGATCAATTTGGCTTGGTTTACAAGGTGCAGCA 600
QY 4031 gcacaatgaatacatcttccagttgtggtcttttcatacaagcaagggcgtgtacc 4090
Db 601 GCACAATGGAATCATCTTCCAGCTTTGTGGCTTTTCATCAACAGCAAGGCGGTGTACC 660
QY 4091 tgaatccctgcagctggagggcagatatgacctgctgattccccagtttagaggccaca 4150
Db 661 TGAATTCCTCGCAGCTGGAGGAGCAGATATGACCTGCTGATTTCCCGAGTTAGAGGCCACA 720
QY 4151 agctctggggccagttccccactgcccattgggggtcagcatagctatagacaagatatctgc 4210

Db 721 AGCTCTGGGGCCAGTTCCCACTGCCATTTGGGTTCAGCATAGCTATAGACAAGATATCTGC 780
QY 4211 tctgtctcctcaacatggaggaatctgttacataagctcttctgtgacctcctcgtgttaag 4270
Db 781 TGTCTGTCTTCAACATGGAGGAATCTGTACATAAAGCTCTTGTGTGACCTCTCTGTTGTAAG 840
QY 4271 tgttgtcagatgtctatgtccagggccatcaacccacagagaactctdgcagcagg 4330
Db 841 TGTGTGGCAGATGTCTATGTCCAGGGCCATCAACCTTAACCCAGAAACTCTGGACAGCAGG 900
QY 4331 catcacagcagaaatcatgtacgtactgtcacagtcacagtcacaggaaggaatcacaagaatctg 4390
Db 901 CATCACAGCAGAAATCATGTACGACTGTGTACAGTGTCCCAAGAGGAATTTACAAGATGACTG 960
QY 4391 cagacatcatgaatcacctatgtggtcccttctcgtgataaaggaagccatgtcaaa 4450
Db 961 CAGACATCATGAATCACCTATGTGGCCCTTGTCTCGGATAAAGAGGAGGCGCATGTCAA 1020
QY 4451 g 4451
Db 1021 G 1021
AC012138 164297 bp DNA HTG 07-JUL-2000
Homo sapiens clone RP11-10K15, WORKING DRAFT SEQUENCE, 34 unordered
pieces
AC012138
AC012138.4 GI:8954133
HTG; HTGS_PHASE1; HTGS_DRAFT.
human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 164297)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArillano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Gallagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Melarim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (20-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 7, 2000 this sequence version replaced GI:7230052.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L3086
Center clone name: 10_K_15
----- Summary Statistics
Sequencing vector: M13; M77815; 98% of reads
Sequencing vector: Plasmid; n/a; 0.0.f% of reads

TITLE
JOURNAL
COMMENT

1.91489361702128Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731
Consensus quality: 142640 bases at least Q40
Consensus quality: 152432 bases at least Q30
Consensus quality: 157096 bases at least Q20
Insert size: 153000; agarose-fp
Insert size: 160997; sum-of-ctnigs
Quality coverage: 3.4 in Q20 bases; agarose-fp
Quality cov.

NOTE: This is a 'working draft' sequence. It currently consists of 34 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1064: contig of 1064 bp in length
1065 1164: gap of 100 bp
1165 2366: contig of 1202 bp in length
2367 2466: gap of 100 bp
2467 3540: contig of 1074 bp in length
3541 3640: gap of 100 bp
3641 4861: contig of 1221 bp in length
4862 4961: gap of 100 bp
4962 5968: contig of 1007 bp in length
5969 6068: gap of 100 bp
6069 7242: contig of 1174 bp in length
7243 7342: gap of 100 bp
7343 8453: contig of 1111 bp in length
8454 8553: gap of 100 bp
8554 9820: contig of 1267 bp in length
9821 9920: gap of 100 bp
9921 10938: contig of 1018 bp in length
10939 11038: gap of 100 bp
11039 12256: contig of 1218 bp in length
12257 12356: gap of 100 bp
12357 14248: contig of 1892 bp in length
14249 14348: gap of 100 bp
14349 15680: contig of 1332 bp in length
15681 15780: gap of 100 bp
15781 17240: contig of 1460 bp in length
17241 17340: gap of 100 bp
17341 18693: contig of 1353 bp in length
18694 18793: gap of 100 bp
18794 20812: contig of 2019 bp in length
20813 20912: gap of 100 bp
20913 24906: contig of 3994 bp in length
24907 25006: gap of 100 bp
25007 28926: contig of 3920 bp in length
28927 29026: gap of 100 bp
29027 34159: contig of 5133 bp in length
34160 34259: gap of 100 bp
34260 37975: contig of 3716 bp in length
37976 38075: gap of 100 bp
38076 43955: contig of 5880 bp in length
43956 44053: gap of 100 bp
44054 48039: contig of 3984 bp in length
48040 48139: gap of 100 bp
48140 53767: contig of 5628 bp in length
53768 53867: gap of 100 bp
53868 59812: contig of 5945 bp in length
59813 59912: gap of 100 bp
59913 63068: contig of 3154 bp in length
63069 63166: gap of 100 bp
63167 70627: contig of 7461 bp in length
70628 70727: gap of 100 bp
70728 77701: contig of 6974 bp in length
77702 77801: gap of 100 bp
77802 85553: contig of 7752 bp in length
85554 85653: gap of 100 bp
85654 94602: contig of 8949 bp in length

* 94603 94702: gap of 100 bp
* 94703 104093: contig of 9391 bp in length
* 104094 104193: gap of 100 bp
* 104194 112395: contig of 8202 bp in length
* 112396 112495: gap of 100 bp
* 112496 123244: contig of 10749 bp in length
* 123245 123344: gap of 100 bp
* 123345 135427: contig of 12083 bp in length
* 135428 135527: gap of 100 bp
* 135528 147434: contig of 11907 bp in length
* 147435 147534: gap of 100 bp
* 147535 164297: contig of 16763 bp in length.
FEATURES
source
1. .164297
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="RPC1-11 Human Male BAC"
1. .1064
/note="assembly_fragment"
1165. .2366
/note="assembly_fragment"
2467. .3540
/note="assembly_fragment"
3641. .4861
/note="assembly_fragment"
4962. .5968
/note="assembly_fragment"
6069. .7242
/note="assembly_fragment"
7343. .8453
/note="assembly_fragment"
8554. .9820
/note="assembly_fragment"
9921. .10938
/note="assembly_fragment"
11039. .12256
/note="assembly_fragment"
12357. .14248
/note="assembly_fragment"
14349. .15680
/note="assembly_fragment"
15781. .17240
/note="assembly_fragment"
17341. .18693
/note="assembly_fragment"
18794. .20812
/note="assembly_fragment"
20913. .24906
/note="assembly_fragment"
25007. .28926
/note="assembly_fragment"
29027. .34159
/note="assembly_fragment"
34260. .37975
/note="assembly_fragment"
clone_end.T7
vector_side:left"
38076. .43955
/note="assembly_fragment"
44056. .48039
/note="assembly_fragment"
48140. .53767
/note="assembly_fragment"
53868. .59812
/note="assembly_fragment"
59913. .63066
/note="assembly_fragment"
clone_end.SP6
vector_side:left"
63167. .70627
/note="assembly_fragment"
70728. .77701

Mon Oct 22 08:29:29 2001

Submitted (07-MAR-2000) Multimegabase Sequencing Center, University of Washington, PO BOX 357730, Seattle, WA 98195, USA
On Apr 13, 2001 this sequence version replaced gi:10305066.

JOURNAL

COMMENT

Center: Multimegabase Sequencing Center
Center code: UWMSC
Web site: http://chroma.mbt.washington.edu/msg_www
Contact: leeroewensystemsbiology.org

Sequencing vector: pUC18; L08752
Chemistry: Dye-terminator Big Dye; 90% of reads
Chemistry: Dye-primer Big Dye; 10% of reads
Assembly program: Phrap; version 0.990399
Insert size: 17200; agarose-fp
Quality coverage: 12.6x in Q20 bases; sum-of-contigs

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the Genbank flat file format but are available as part of this entry's ASN.1 file.

* NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 173364: contig of 173364 bp in length.

FEATURES

source

1. 173364
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="15"
/map="15q14"
/clone="RP11-43D14"
/clone_lib="RPCI human BAC library 11"
/notes="This clone overlaps RP11-325N19 and RP11-521C20"

BASE COUNT
ORIGIN

Query Match 9.8%; Score 539; DB 69; Length 173364;
Best Local Similarity 100.0%; Pred. No. 2.2e-292;

Matches 539; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1077 cagattcaaggaacagaaacagaaattcaactcactggtgataaattgagccatccaaatgta 1136

Db 72963 CAGATTCAAGGAACAGAAACAGAAATTCAACTCACTGCTGTAATAATTGAGCCATCCAAATGTA 72904

QY 1137 gtacgtaccttgcattgaatctcaagagcaagcaactccatcggtgacatttta 1196

Db 72903 GTACGCTACCTTGCAATGAATCTCAAGAGCAAGACACATCCATCGGTGGACATTTTA 72844

QY 1197 gggagacattagtggtggtctcttctgctgcacacctgagccactcagcccccctccct 1256

Db 72843 GTGGAGCACATTAGTGGGGTCTCTCTGCTGCACACCTGAGCCACTCAGGCCCATCCCT 72784

QY 1257 gtgcacagcttcgcaggtacacagctcagctcctgtcagcccttgattatctgcacagc 1316

Db 72783 GTGCATCAGCTTCGCGAGGTACACAGCTCAGCTCCTGTGTCAGGCCCTTGATTATCTGCACAGC 72724

QY 1317 aattctgtgtgcataaagctcctgagtcacatctaatgtctgtgtgagtcagagggcacc 1376

Db 72723 AATTCTGTGTGCATAAAGGTCCTGTAGTGCATCTAATGTCTTGGTGGATGTCAGAGGCACC 72664

misc_feature /note="assembly_fragment"
77802..85553
misc_feature /note="assembly_fragment"
85654..94602
misc_feature /note="assembly_fragment"
94703..104093
/note="assembly_fragment"

Query Match 9.8%; Score 539; DB 62; Length 164297;
Best Local Similarity 100.0%; Pred. No. 2.2e-292;
Matches 539; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1077 cagattcaaggaacagaaacagaaattcaactcactggtgataaattgagccatccaaatgta 1136
Db 60669 CAGATTCAAGGAACAGAAACAGAAATTCAACTCACTGCTGTAATAATTGAGCCATCCAAATGTA 60610
QY 1137 gtacgtaccttgcattgaatctcaagagcaagcaactccatcggtgacatttta 1196
Db 60609 GTACGCTACCTTGCAATGAATCTCAAGAGCAAGACGACTCCATCTGCTGGACATTTTA 60550
QY 1197 gtgcagcacattagtggtggtctcttctgctgcacacctgagccactcagcccccctccct 1256
Db 60549 GTGCAGCACATTAGTGGGCTCTCTCTGCTGCACACCTGAGCCACTCAGGCCCATCCCT 60490
QY 1257 gtgcacagcttcgcaggtacacagctcagctcctctcagcccttgattatctgcacagc 1316
Db 60489 GTGCATCAGCTTCGCGAGGTACACAGCTCAGCTCCTGTGTCAGGCCCTTGATTATCTGCACAGC 60430
QY 1317 aattctgtgtgcataaagctcctgagtcacatctaatgtcttctgtgagtcagaaagggcacc 1376
Db 60429 AATTCTGTGTGCATAAAGGTCCTGAGTGCATCTAATGTCTTGGTGGATGCAGAGGCACC 60370
QY 1377 gtcaagattacggactatagattcttaagcgcctcgcagacatttgcaaggaggatgtg 1436
Db 60369 GTCAAGATTACGGACTATAGCATTTCTTAAGCGCTCGCAGACATTTGCAAGGAGGATGTG 60310
QY 1437 tttagcaaacccaggtctgttagtgacaatgctctgctctataaacggggaagaaa 1496
Db 60309 TTTGAGCAAAACCGAGTTCGTTTTAGTGCACAAATGCTGCTTATAAAGGGGGAAGAAA 60250
QY 1497 ggagatgttggcgtcttggcctctctgctgctcctcgcacagcaaggaaggaatgtga 1556
Db 60249 GGAGATGTTGGCGCTTGGCGCTTCTGCTGCTGCTCCTCGCCAGCAAGGACAGGAATGTGA 60190
QY 1557 gagtcacctgtgacacatcccttagtgaacttaccagctacattcaagatttctaaagaa 1615
Db 60189 GAGTACCTGTGACCATCCCTAGTGAATTTACCAAGTCTTACCAAGTCTTCAAGATTTTCTTAAGAA 60131

RESULT 9

AC025168/c 173364 bp DNA HTG 13-APR-2001
LOCUS Homo sapiens chromosome 15 clone RP11-43D14 map 15q14, ***
DEFINITION SEQUENCING IN PROGRESS ***, 1 ordered pieces.
AC025168
VERSION AC025168.6 GI:13621223
KEYWORDS HTG; HTGS_PHASE2; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 173364)
AUTHORS Rowen, L., Madan, A., Qin, S., Baradarani, L., Birditt, B.,
Burke, J., Dors, M., Fleetwood, P., Kaur, A., Madan, A., Nesbitt, R.,
Pate, D. and Hood, L.
Sequencing of human chromosome 15 D15S146-D15S117 region
Unpublished
TITLE 2 (bases 1 to 173364)
JOURNAL
AUTHORS Rowen, L., Madan, A., Qin, S., Abbasi, N., Baradarani, L., Birditt, B.,
Bloom, S., Dors, M., Dickhoff, R., Fleetwood, P., Harrison, G.,
James, R., Kaur, A., Madan, A., Owen, M.P., Ratcliffe, A., Shaffer, T.
and Hood, L.
Direct Submission
TITLE

```

QY 1377 gtaagattacgagcattcttaagcgccctcgagacatttcgaagaggatgtg 1436
Db 7263 CTCAGATTACGGACTATAGCATTTCTAAGCGCTCGCAGACATTTCAAGGAGGATGTG 72604
QY 1437 tttaagcaaacccgagtgcttttagtgacaaatgctctgcttataaacggggaagaa 1496
Db 72603 TTTGAGCAAAACCCGAGTTCGTTTTAGTGACAATGCTCTGCTTATAAACGGGGAAGAA 72544
QY 1497 ggagatgtttggcgtcttgccctctgctgctccctcagcaagacaggaatgtgga 1556
Db 72543 GGAGATGTTGGCGCTTGGCGCTTCTGCTGCTGCTCGCCCTCAGCCAAAGACAGGAATGTGGA 72484
QY 1557 gactaccctgagcaccctcctagtgactaccagctgactttcgaagattttctaaagaa 1615
Db 72483 GAGTACCCTGTGACCATCCCTAGTGACTTACCAGCTGACTTTCAAGATTTTCTAAGAA 72425

RESULT 10
AC021755 157900 bp DNA HTG 21-MAR-2001
DEFINITION Homo sapiens chromosome 15 clone RP11-521C20 map 15q14, WORKING
DRAFT SEQUENCE, 6 ordered pieces.
ACCESSION AC021755
VERSION AC021755.6 GI:13399357
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE human.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 157900)
AUTHORS Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G., Kaur,A., Madan,A.,
Nesbitt,R., Traicoff,R. and Hood,L.
Sequencing of human chromosome 15 D15S146-D15S117 region
Unpublished
2 (bases 1 to 157900)
AUTHORS Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
Bloom,S., Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G.,
James,R., Kaur,A., Madan,A., Owen,M.P., Ratcliffe,A., Shaffer,T.
and Hood,L.
Direct Submission
Submitted (20-JAN-2000) Multimegabase Sequencing Center, University
of Washington, PO BOX 357730, Seattle, WA 98195, USA
On Mar 21, 2001 this sequence version replaced gi:8272664.
-----
Center: Multimegabase Sequencing Center
Genome Center
Center code: UWMS
Web site: http://chroma.mbt.washington.edu/msg_www
Contact: leetowen@systemsbiology.org
-----
Sequencing vector: pUC18; L08752
Chemistry: Dye-terminator Big Dye; 90% of reads
Chemistry: Dye-primer Big Dye; 10% of reads
Assembly program: Phrap; version 0.990399
Insert size: 150000; agarose-fp
Quality coverage: 10.2x in Q20 bases; sum-of-contigs
-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.

```

```

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 67961: contig of 67961 bp in length
* 67962 68061: gap of unknown length
* 68062 76767: contig of 8706 bp in length
* 76768 76867: gap of unknown length
* 76868 79623: contig of 2756 bp in length
* 79624 90362: contig of 10639 bp in length
* 90363 90462: gap of unknown length
* 90463 93020: contig of 2558 bp in length
* 93021 93120: gap of unknown length
* 93121 157900: contig of 64780 bp in length.
FEATURES
Location/Qualifiers
1..157900
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="15"
/map="15q14"
/clone="RP11-521C20"
/clone_lib="RPCI human BAC library 11"
/note="This clone overlaps RP11-43D14 and CTD-2006D8"
BASE COUNT 43137 a 35607 c 36148 g 42500 t 508 others
ORIGIN
Query Match 9.4%; Score 519; DB 66; Length 157900;
Best Local Similarity 99.8%; Pred. No. 4.4e-281;
Matches 569; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4955 ggtgtctgtgctatttctgtacagctatagagtgactactacagaatcttatttaacc 5014
Db 52726 GGTGTCTGTGCTATTCTGTACAGCTATAGAGTACTACTACAGAATCTTATTTAACC 52785
QY 5015 cttaagaaactgcgttaacctattcaaacagacagagcttatactggaataatggaa 5074
Db 52786 CTAAAGAACTGCTGTACCTCATTTCAACACAGACAGAGGCTTACTCGATATATGGAA 52845
QY 5075 gttgtacattatcatataatttaaataattcttaagaagaggtcgtgcagtggtcca 5134
Db 52846 GTTGTACATTCATCATAATTTAAATTTAAATTTCTAAGAAGAGGCTGGTGCAGTGGCTCA 52905
QY 5135 cacccttaatccagcactttggaaagccaaagcaggaagagactcttgaaaccaggaggt 5194
Db 52906 CACCTTTAATCCAGCACCCTTTGGGAAGCCAAAGCGGAGGACTGCTTGAACACAGGAGTT 52965
QY 5195 tgagaccagcctgagcaacaaagcaagccccctctctataaaactaaaaaataatagtt 5254
Db 52966 TGAGACCAGCTGAGCAACAAAGCAAGACCCCATCTCTATAAANAATAAATAATAGTT 53025
QY 5255 gggcatggtggcacaatgctctagtcacagctatccagaggtgagatgagatcatctga 5314
Db 53026 GGGCATGGTGGCACATGCTGTAGTCCAGCTACTCCAGAGGCTGAGATGATCATCTGA 53085
QY 5315 gcctcagagaggttgaggtcagtgagctgactgcacactgcactgcactgcagctggagc 5374
Db 53086 GCCTCAGAGGTTGAGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 53145
QY 5375 aacagagcaagaccctgtcttaaaaaaagaaaaaataatcttttcttaagaagct 5434
Db 53146 AACAGAGCAAGACCCCTGTCTTAAAAAAGAAAAAATAATTTTCTTCTAAGAGCT 53205
QY 5435 gtccatacaaggttagctgtttagttttagtttttcctatgtgtaataataataatttatcttt 5494
Db 53206 GTCCTCAAAAGTTGAGCTTTGTTAGTTTTCATGTTGTAATATATATAAATTTATCTTTT 53265
QY 5495 gggataataaataatgctttcatatcacctgc 5524
Db 53266 GGGATATAATAAATGCTTTTCATATACCTGC 53295

```

RESULT 11

```

AC090997      61029 bp      DNA      HTG      22-MAR-2001
LOCUS      Homo sapiens chromosome 15 clone RP11-521C20 map 15, LOW-PASS
DEFINITION      SEQUENCE SAMPLING.
AC090997      61029 bp      DNA      HTG      22-MAR-2001
VERSION      AC090997.1 GI:13431057
KEYWORDS      HTG; HTGS_PHASE0.
SOURCE      human.
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
              1 (bases 1 to 61029)
AUTHORS      Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE      Homo sapiens chromosome 15, clone RP11-521C20
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 61029)
              Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
              Barne, N., Bastien, V., Boguslavsky, L., Bouckhalter, B., Brown, A.,
              Camarata, J., Campopiano, A., Chang, J., Choepel, Y., Colangelo, M.,
              Collins, S., Collamore, A., Cooke, P., DeArelano, K., Dewar, K.,
              Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D.,
              Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
              Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W.,
              Iliev, I., Johnson, R., Jones, C., Karatas, A., LaRoque, K.,
              Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
              MacLean, C., Macdonald, P., Marquis, N., Matthews, C., McCarthy, M.,
              McEwan, P., McKernan, K., McPheters, R., Meldrum, J., Meneus, L.,
              Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C.,
              Norman, K., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
              Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
              Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
              Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S.,
              Severy, P., Sougne, C., Spencer, B., Stange-Thomann, N.,
              Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S.,
              Theodore, J., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
              Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
              Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
              Direct Submission
              Submitted (22-MAR-2001) Whitehead Institute/MIT Center for Genome
              Research, 320 Charles Street, Cambridge, MA 02141, USA
              All repeats were identified using RepeatMasker:
              Smit, A.F.A. & Green, P. (1996-1997)
              http://ftp.genome.washington.edu/RM/RepeatMasker.html
              ----- Genome Center
              Center: Whitehead Institute/ MIT Center for Genome Research
              Center code: WIBR
              Web site: http://www-seq.wi.mit.edu
              Contact: sequence_submissions@genome.wi.mit.edu
              ----- Project Information
              Center project name: L13141
              Center clone name: 521_C-20
              -----
              * NOTE: This record contains 82 individual
              * sequencing reads that have not been assembled into
              * contigs. Runs of N are used to separate the reads
              * and the order in which they appear is completely
              * arbitrary. Low-pass sequence sampling is useful for
              * identifying clones that may be gene-rich and allows
              * overlap relationships among clones to be deduced.
              * However, it should not be assumed that this clone
              * will be sequenced to completion. In the event that
              * the record is updated, the accession number will
              * be preserved.
              *
              * 571 670: gap of 100 bp
              * 671 1297: contig of 627 bp in length
              * 1298 1397: gap of 100 bp
              * 1398 2087: contig of 690 bp in length
              * 2088 2187: gap of 100 bp
              * 2188 2956: contig of 769 bp in length
              * 2957 3056: gap of 100 bp
              * 3057 3690: contig of 634 bp in length
              * 3691 3790: gap of 100 bp
              *
              * 3791 4405: contig of 615 bp in length
              * 4406 4505: gap of 100 bp
              * 4506 5161: contig of 656 bp in length
              * 5162 5261: gap of 100 bp
              * 5262 5867: contig of 606 bp in length
              * 5868 5967: gap of 100 bp
              * 5968 6630: contig of 663 bp in length
              * 6631 6730: gap of 100 bp
              * 6731 7361: contig of 631 bp in length
              * 7362 7461: gap of 100 bp
              * 7462 8083: contig of 622 bp in length
              * 8084 8183: gap of 100 bp
              * 8184 8853: contig of 670 bp in length
              * 8854 8953: gap of 100 bp
              * 8954 9677: contig of 724 bp in length
              * 9678 9777: gap of 100 bp
              * 9778 10498: contig of 721 bp in length
              * 10499 10598: gap of 100 bp
              * 10599 11211: contig of 613 bp in length
              * 11212 11311: gap of 100 bp
              * 11312 11946: contig of 635 bp in length
              * 11947 12046: gap of 100 bp
              * 12047 12682: contig of 636 bp in length
              * 12683 12782: gap of 100 bp
              * 12783 13443: contig of 661 bp in length
              * 13444 13543: gap of 100 bp
              * 13544 14212: contig of 669 bp in length
              * 14213 14312: gap of 100 bp
              * 14313 14919: contig of 607 bp in length
              * 14920 15019: gap of 100 bp
              * 15020 15680: contig of 661 bp in length
              * 15681 15780: gap of 100 bp
              * 15781 16337: contig of 557 bp in length
              * 16338 16437: gap of 100 bp
              * 16438 17112: contig of 675 bp in length
              * 17113 17212: gap of 100 bp
              * 17213 17901: contig of 689 bp in length
              * 17902 18001: gap of 100 bp
              * 18002 18749: contig of 748 bp in length
              * 18750 18849: gap of 100 bp
              * 18850 19453: contig of 604 bp in length
              * 19454 19553: gap of 100 bp
              * 19554 20189: contig of 636 bp in length
              * 20190 20289: gap of 100 bp
              * 20290 20942: contig of 653 bp in length
              * 20943 21042: gap of 100 bp
              * 21043 21648: contig of 606 bp in length
              * 21649 21748: gap of 100 bp
              * 21749 22357: contig of 609 bp in length
              * 22358 22457: gap of 100 bp
              * 22458 23104: contig of 647 bp in length
              * 23105 23204: gap of 100 bp
              * 23205 23840: contig of 636 bp in length
              * 23841 23940: gap of 100 bp
              * 23941 24516: contig of 576 bp in length
              * 24517 24616: gap of 100 bp
              * 24617 25314: contig of 698 bp in length
              * 25315 25414: gap of 100 bp
              * 25415 26037: contig of 623 bp in length
              * 26038 26137: gap of 100 bp
              * 26138 26749: contig of 612 bp in length
              * 26750 26849: gap of 100 bp
              * 26850 27497: contig of 648 bp in length
              * 27498 27597: gap of 100 bp
              * 27598 28259: contig of 662 bp in length
              * 28260 28359: gap of 100 bp
              * 28360 29041: contig of 682 bp in length
              * 29042 29141: gap of 100 bp
              * 29142 29807: contig of 666 bp in length
              * 29808 29907: gap of 100 bp
              * 29908 30480: contig of 573 bp in length
              * 30481 30580: gap of 100 bp
              * 30581 31289: contig of 709 bp in length
              *

```

	*	31290	31389:	gap of	100 bp	
	*	31390	32097:	contig of 708 bp	in length	
	*	32098	32197:	gap of	100 bp	
	*	32198	32914:	contig of 717 bp	in length	
	*	32915	33014:	gap of	100 bp	
	*	33015	33618:	contig of 604 bp	in length	
	*	33619	33718:	gap of	100 bp	
	*	33719	34358:	contig of 640 bp	in length	
	*	34359	34458:	gap of	100 bp	
	*	34459	34977:	contig of 519 bp	in length	
	*	34978	35077:	gap of	100 bp	
	*	35078	35689:	contig of 612 bp	in length	
	*	35690	35789:	gap of	100 bp	
	*	35790	36386:	contig of 597 bp	in length	
	*	36387	36486:	gap of	100 bp	
	*	36487	37092:	contig of 606 bp	in length	
	*	37093	37192:	gap of	100 bp	
	*	37193	37864:	contig of 672 bp	in length	
	*	37865	37964:	gap of	100 bp	
	*	37965	38626:	contig of 662 bp	in length	
	*	38627	38726:	gap of	100 bp	
	*	38727	39348:	contig of 622 bp	in length	
	*	39349	39448:	gap of	100 bp	
	*	39449	40072:	contig of 624 bp	in length	
	*	40073	40172:	gap of	100 bp	
	*	40173	40882:	contig of 710 bp	in length	
	*	40883	40982:	gap of	100 bp	
	*	40983	41705:	contig of 723 bp	in length	
	*	41706	41805:	gap of	100 bp	
	*	41806	42412:	contig of 607 bp	in length	
	*	42413	42512:	gap of	100 bp	
	*	42513	43158:	contig of 646 bp	in length	
	*	43159	43258:	gap of	100 bp	
	*	43259	43880:	contig of 622 bp	in length	
	*	43881	43980:	gap of	100 bp	
	*	43981	44618:	contig of 638 bp	in length	
	*	44619	44718:	gap of	100 bp	
	*	44719	45359:	contig of 641 bp	in length	
	*	45360	45459:	gap of	100 bp	
	*	45460	46088:	contig of 629 bp	in length	
	*	46089	46188:	gap of	100 bp	
	*	46189	46855:	contig of 667 bp	in length	
	*	46856	46955:	gap of	100 bp	
	*	46956	47642:	contig of 687 bp	in length	
	*	47643	47742:	gap of	100 bp	
	*	47743	48295:	contig of 553 bp	in length	
	*	48296	48395:	gap of	100 bp	
	*	48396	49021:	contig of 626 bp	in length	
	*	49022	49121:	gap of	100 bp	
	*	49122	49825:	contig of 704 bp	in length	
	*	49826	49925:	gap of	100 bp	
	*	49926	50543:	contig of 618 bp	in length	
	*	50544	50643:	gap of	100 bp	
	*	50644	51266:	contig of 623 bp	in length	
	*	51267	51366:	gap of	100 bp	
Query Match			3.8%;	Score 211;	DB 78;	Length 61029;
Best Local Similarity			99.4%;	Pred. No.	5.7e-107;	
Matches 311; Conservative			0;	Mismatches	2;	Indels 0; Gaps
QY	5156	gggaagccaaggcagaagactgttgaaccaggagttagacaccgctgagcaacaa	5215			
Db	47983	GGGAAGCCAAAGCAGGAGACTGCTTCAAACACGAGGATTGAGACCAGCCTGAGCAACA	4804			
QY	5216	aqcaagcccattctataaaaaactaaaaattatgttggcattggtggcacatgcctg	5275			
Db	48043	AGCAAGCCCCATCTCTATAAAAACTATAAAAATTANTTTGGCATGTGGGCATCGCTG	48102			
QY	5276	tagtcccgactactccagagctgagatgatcatctgacctcaggagttgagctgc	5335			
Db	48103	TAGTCCCAGCTACTCCAGAGCCTGAGATGATCATCTGAGCCCTGAGGAGTTGAGGCTG	48162			
QY	5336	adtgagctgtgactgcgcaccctgactccctgctggacaacagacagaccctgtctt	5395			

* will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1 570: contig of 570 bp in length
 571 670: gap of 100 bp
 671 1297: contig of 627 bp in length
 1298 1397: gap of 100 bp
 1398 2087: contig of 690 bp in length
 2088 2187: gap of 100 bp
 2188 2956: contig of 769 bp in length
 2957 3056: gap of 100 bp
 3057 3650: contig of 634 bp in length
 3651 3790: gap of 100 bp
 3791 4405: contig of 615 bp in length
 4406 4505: gap of 100 bp
 4506 5161: contig of 656 bp in length
 5162 5261: gap of 100 bp
 5262 5867: contig of 606 bp in length
 5868 5967: gap of 100 bp
 5968 6630: contig of 663 bp in length
 6631 6730: gap of 100 bp
 6731 7361: contig of 631 bp in length
 7362 7461: gap of 100 bp
 7462 8083: contig of 622 bp in length
 8084 8183: gap of 100 bp
 8184 8853: contig of 670 bp in length
 8854 8953: gap of 100 bp
 8954 9677: contig of 724 bp in length
 9678 9777: gap of 100 bp
 9778 10498: contig of 721 bp in length
 10499 10598: gap of 100 bp
 10599 11211: contig of 613 bp in length
 11212 11311: gap of 100 bp
 11312 11946: contig of 635 bp in length
 11947 12046: gap of 100 bp
 12047 12682: contig of 636 bp in length
 12683 12782: gap of 100 bp
 12783 13443: contig of 661 bp in length
 13444 13543: gap of 100 bp
 13544 14212: contig of 669 bp in length
 14213 14312: gap of 100 bp
 14313 14919: contig of 607 bp in length
 14920 15019: gap of 100 bp
 15020 15680: contig of 661 bp in length
 15681 15780: gap of 100 bp
 15781 16337: contig of 557 bp in length
 16338 16437: gap of 100 bp
 16438 17112: contig of 675 bp in length
 17113 17212: gap of 100 bp
 17213 17901: contig of 689 bp in length
 17902 18001: gap of 100 bp
 18002 18749: contig of 748 bp in length
 18750 18849: gap of 100 bp
 18850 19453: contig of 604 bp in length
 19454 19553: gap of 100 bp
 19554 20189: contig of 636 bp in length
 20190 20289: gap of 100 bp
 20290 20943: contig of 653 bp in length
 20943 21042: gap of 100 bp
 21043 21648: contig of 606 bp in length
 21649 21748: gap of 100 bp
 21749 22357: contig of 609 bp in length
 22358 22457: gap of 100 bp
 22458 23104: contig of 647 bp in length
 23105 23204: gap of 100 bp
 23205 23840: contig of 636 bp in length
 23841 23940: gap of 100 bp
 23941 24516: contig of 576 bp in length
 24517 24616: gap of 100 bp
 24617 25314: contig of 698 bp in length
 25315 25414: gap of 100 bp
 25415 26037: contig of 623 bp in length
 26038 26137: gap of 100 bp

26138 26749: contig of 612 bp in length
 26750 26849: gap of 100 bp
 26850 27497: contig of 648 bp in length
 27498 27597: gap of 100 bp
 27598 28259: contig of 662 bp in length
 28260 28359: gap of 100 bp
 28360 29041: contig of 682 bp in length
 29042 29141: gap of 100 bp
 29142 29807: contig of 666 bp in length
 29808 29907: gap of 100 bp
 29908 30480: contig of 573 bp in length
 30481 30580: gap of 100 bp
 30581 31289: contig of 709 bp in length
 31290 31389: gap of 100 bp
 31390 32097: contig of 708 bp in length
 32098 32197: gap of 100 bp
 32198 32914: contig of 717 bp in length
 32915 33014: gap of 100 bp
 33015 33618: contig of 604 bp in length
 33619 33718: gap of 100 bp
 33719 34358: contig of 640 bp in length
 34359 34458: gap of 100 bp
 34459 34977: contig of 519 bp in length
 34978 35077: gap of 100 bp
 35078 35689: contig of 612 bp in length
 35690 35789: gap of 100 bp
 35790 36386: contig of 597 bp in length
 36387 36486: gap of 100 bp
 36487 37092: contig of 606 bp in length
 37093 37192: gap of 100 bp
 37193 37864: contig of 672 bp in length
 37865 37964: gap of 100 bp
 37965 38626: contig of 662 bp in length
 38627 38726: gap of 100 bp
 38727 39348: contig of 622 bp in length
 39349 39448: gap of 100 bp
 39449 40072: contig of 624 bp in length
 40073 40172: gap of 100 bp
 40173 40882: contig of 710 bp in length
 40883 40982: gap of 100 bp
 40983 41705: contig of 723 bp in length
 41706 41805: gap of 100 bp
 41806 42412: contig of 607 bp in length
 42413 42512: gap of 100 bp
 42513 43158: contig of 646 bp in length
 43159 43258: gap of 100 bp
 43259 43880: contig of 622 bp in length
 43881 43980: gap of 100 bp
 43981 44618: contig of 638 bp in length
 44619 44718: gap of 100 bp
 44719 45359: contig of 641 bp in length
 45360 45459: gap of 100 bp
 45460 46088: contig of 629 bp in length
 46089 46188: gap of 100 bp
 46189 46855: contig of 667 bp in length
 46856 46955: gap of 100 bp
 46956 47642: contig of 687 bp in length
 47643 47742: gap of 100 bp
 47743 48295: contig of 553 bp in length
 48296 48395: gap of 100 bp
 48396 49021: contig of 626 bp in length
 49022 49121: gap of 100 bp
 49122 49825: contig of 704 bp in length
 49826 49925: gap of 100 bp
 49926 50543: contig of 618 bp in length
 50544 50643: gap of 100 bp
 50644 51266: contig of 623 bp in length
 51267 51366: gap of 100 bp

Query Match 3.1%; Score 171; DB 78; Length 61029;
 Best Local Similarity 100.0%; Pred. No. 2.4e-84;
 Matches 171; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 3469 gatactgacataacgctgtgttcacgagcgacgaagttagatcattccatcccaagaac 3528
Db 13127 TATCTGATGATAGACGCTGTTCAGCGCGCGCAAGTTAGATGATTCATCCCAAGAAC 13068
QY 3529 ttctggagtgcatttgattgatttgcactttacacacacagctttctgcccactgctg 3588
Db 13067 TTTCTGGAGTGTGCAATTTGATATTTGCTCACTTCTACCAACACAGCTTTCTGCCACTGCTG 13008
QY 3589 aaattatcacactatcatataatcatcccaagagtttccagacttcagg 3639
Db 13007 AAATTATCTACACTATCATGAATCATCCCAAGAGTTTCCAGCACTTCAGG 12957

RESULT 13
AC012377
LOCUS AC012377 197610 bp DNA HTG 13-APR-2001
DEFINITION Homo sapiens chromosome 15 clone RP11-325N19 map 15q14, ***
SEQUENCING IN PROGRESS ***, 1 ordered pieces.
AC012377
VERSION AC012377.4 GI:13621218
KEYWORDS HTG: HTGS_PHASE2; HTGS_ACTIVEFIN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 197610)
Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B.,
Bloom,S., Dors,M., Dickhoff,R., Harrison,G., James,K., Madan,A.,
Owen,M.P., Ratcliffe,A., Shaffer,T. and Hood,L.
Direct Submission
Submitted (26-OCT-1999), Multimegabase Sequencing Center, University
of Washington, PO BOX 357730, Seattle, WA 98195, USA
On Apr 13, 2001 this sequence version replaced gi:10567837.
----- Genome Center Sequencing Center
Center: Multimegabase Sequencing Center
Center code: UMSC
Web site: http://chroma.mbt.washington.edu/msg_www
Contact: leerowensystemsbiology.org
----- Summary Statistics
Sequencing vector: pUC18; L08752
Chemistry: Dye-terminator Big Dye; 90% of reads
Assembly: Dye-primer Big Dye; 10% of reads
Assembly program: Phrap; version 0.990399
Insert size: 197000; agarose-fp
Quality coverage: 13x in Q20 bases; sum-of-contigs
-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
Genbank flat file format but are available as part
of this entry's ASN.1 file.
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as runs of N. However the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 197610: contig of 197610 bp in length.
* Location/Qualifiers

```

source

```

1. .197610
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="15"
/map="15q14"
/clone="RP11-325N19"
/clone_lib="human RPI BAC library 11"
/note="This clone overlaps RP11-37C7 and RP11-43D14"
BASE COUNT 61560 a 41034 c 39114 g 55902 t
ORIGIN

```

```

Query Match 2.8%; Score 155; DB 62; Length 197610;
Best Local Similarity 100.0%; Pred. NO. 3e-75;
Matches 155; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 421 agtgatgatttgaactggcttaccacgtgcagtcatttctcagcagcatacaaac 480
Db 188078 AGTGATGATCTTTGAACCTGGCTTACCAGTCGACGTCATTTCTCAGCAGCATAACAAG 188137
QY 481 cccctcccaagtcttttcatgaagaatgctggaagcggtcagcagcagcagcaga 540
Db 188138 CCCTCCCAAGTCTTTTCATGAAGAAATGCTGAAAGCGGCTCAGGAGGAGCAGCAGA 188197
QY 541 ggctgttgagcccaagcggaaagagcagcagag 575
Db 188198 GGCTGTGGAGCGCAAGCGGAAAGAGCAGCAGAG 188232

```

RESULT 14

```

AC012138 164297 bp DNA HTG 07-JUL-2000
LOCUS Homo sapiens clone RP11-10K15, WORKING DRAFT SEQUENCE, 34 unordered
DEFINITION pieces.
AC012138
VERSION AC012138.4 GI:8954133
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 164297)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-10K15
Unpublished
2 (bases 1 to 164297)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boquslavkiy,L., Boukhalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collamore,A.,
Cooke,P., Dearrellano,K., Dewar,K., Domino,M., Donegan,L., Doyle,M.,
Ferrelira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Savery,P.,
Stange-Thoman,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye.W.J., Zimmer,A. and Zody,M.
Direct Submission

```

TITLE

```

JOURNAL Submitted (20-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Jul 7, 2000 this sequence version replaced gi:7230052.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center

```

```

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information

```

FEATURES

Center project name: L3086
 Center clone name: 10_K15
 ----- Summary Statistics
 Sequencing vector: M13; M7815; 98% of reads
 Sequencing vector: Plasmid; n/a; %0.f% of reads
 1.91489361702128Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731
 Consensus quality: 142640 bases at least Q40
 Consensus quality: 152432 bases at least Q30
 Consensus quality: 157096 bases at least Q20
 Insert size: 153000; agarose-fp
 Insert size: 160997; sum-of-contigs
 Quality coverage: 3.4 in Q20 bases; agarose-fp
 Quality cov.

NOTE: This is a 'working draft' sequence. It currently consists of 34 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1 1064: contig of 1064 bp in length
1065 1164: gap of 100 bp
1165 2366: contig of 1202 bp in length
2367 2466: gap of 100 bp
2467 3540: contig of 1074 bp in length
3541 3640: gap of 100 bp
3641 4861: contig of 1221 bp in length
4862 4961: gap of 100 bp
4962 5968: contig of 1007 bp in length
5969 6068: gap of 100 bp
6069 7242: contig of 1174 bp in length
7243 7342: gap of 100 bp
7343 8453: contig of 1111 bp in length
8454 8553: gap of 100 bp
8554 9820: contig of 1267 bp in length
9821 9920: gap of 100 bp
9921 10938: contig of 1018 bp in length
10939 11038: gap of 100 bp
11039 12256: contig of 1218 bp in length
12257 12356: gap of 100 bp
12357 14248: contig of 1892 bp in length
14249 14348: gap of 100 bp
14349 15680: contig of 1332 bp in length
15681 15780: gap of 100 bp
15781 17240: contig of 1460 bp in length
17241 17340: gap of 100 bp
17341 18693: contig of 1353 bp in length
18694 18793: gap of 100 bp
18794 20812: contig of 2019 bp in length
20813 20912: gap of 100 bp
20913 24906: contig of 3994 bp in length
24907 25006: gap of 100 bp
25007 28926: contig of 3920 bp in length
28927 29026: gap of 100 bp
29027 34159: contig of 5133 bp in length
34160 34259: gap of 100 bp
34260 37975: contig of 3716 bp in length
37976 38075: gap of 100 bp
38076 43955: contig of 5880 bp in length
43956 44055: gap of 100 bp
44056 48039: contig of 3984 bp in length
48040 48139: gap of 100 bp
48140 53767: contig of 5628 bp in length
53768 53867: gap of 100 bp
53868 59812: contig of 5945 bp in length
59813 59912: gap of 100 bp
59913 63066: contig of 3154 bp in length
63067 63166: gap of 100 bp
63167 70627: contig of 7461 bp in length
70628 70727: gap of 100 bp

```

FEATURES

source

```

* 70728 77701: contig of 6974 bp in length
* 77702 77801: gap of 100 bp
* 77802 85553: contig of 7752 bp in length
* 85554 85653: gap of 100 bp
* 85654 94602: contig of 8949 bp in length
* 94603 94702: gap of 100 bp
* 94703 104093: contig of 9391 bp in length
* 104094 104193: gap of 100 bp
* 104194 112395: contig of 8202 bp in length
* 112396 112495: gap of 100 bp
* 112496 123244: contig of 10749 bp in length
* 123245 123344: gap of 100 bp
* 123345 135427: contig of 12083 bp in length
* 135428 135527: gap of 100 bp
* 135528 147434: contig of 11907 bp in length
* 147435 147534: gap of 100 bp
* 147535 164297: contig of 16763 bp in length.
Location/Qualifiers
1. 164297
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-10K15"
/clone_lib="RPC1-11 Human Male BAC"
1. 1064
/note="assembly_fragment"
1165. 2366
/note="assembly_fragment"
2467. 3540
/note="assembly_fragment"
3641. 4861
/note="assembly_fragment"
4962. 5968
/note="assembly_fragment"
6069. 7242
/note="assembly_fragment"
7343. 8453
/note="assembly_fragment"
8554. 9820
/note="assembly_fragment"
9921. 10938
/note="assembly_fragment"
11039. 12256
/note="assembly_fragment"
12357. 14248
/note="assembly_fragment"
14349. 15680
/note="assembly_fragment"
15781. 17240
/note="assembly_fragment"
17341. 18693
/note="assembly_fragment"
18794. 20812
/note="assembly_fragment"
20913. 24906
/note="assembly_fragment"
25007. 28926
/note="assembly_fragment"
29027. 34159
/note="assembly_fragment"
34260. 37975
/note="assembly_fragment"
clone_end:77
vector_side:left
38076. 43955
/note="assembly_fragment"
44056. 48039
/note="assembly_fragment"
48140. 53767
/note="assembly_fragment"
53868. 59812
/note="assembly_fragment"
59913. 63066
/note="assembly_fragment"

```


